

La edición digital de textos literarios: planteamientos y perspectivas de futuro

Digital Scholarly Editions for: Some Approaches and Future Perspectives

ELENA GONZÁLEZ-BLANCO

Departamento de Literatura Española y Teoría de la Literatura
Facultad de Filología. UNED
Paseo Senda del Rey 7. Madrid, 28040
egonzalezblanco@flog.uned.es

RECIBIDO: 31 DE ENERO DE 2014
ACEPTADO: 24 DE ABRIL DE 2014

Resumen: La edición digital de textos literarios se encuentra sumida en un proceso de evolución constante con amplísimas perspectivas de desarrollo gracias a las soluciones tecnológicas disponibles. Sin embargo, las posibilidades teóricas de la actual sociedad de la información contrastan ampliamente con el panorama real de edición que nos encontramos en España, donde la llamada “brecha digital” se hace patente. Este trabajo se ocupa, en primer lugar, de evaluar el panorama internacional de recursos, herramientas y proyectos en el ámbito de la edición digital de textos literarios para después analizar sus claves de evolución y otros aspectos relevantes, como el acceso a los contenidos en abierto, el uso de interfaces amigables y accesibles, el cumplimiento de los estándares, el trabajo interdisciplinar y la interoperabilidad. El objetivo final de estos análisis y reflexiones es ofrecer nuevas pautas y perspectivas de futuro que trasciendan más allá del texto como mero objeto de estudio.

Abstract: Digital scholarly edition is conditioned by a continuous and challenging process of change due to the availability of many different technical solutions. However, those theoretical possibilities face with the real situation of editions in Spain, where the so-called “digital gap” grows constantly. This paper deals with the analysis of the international panorama of resources, tools and projects related to digital scholarly editions. Its objective is to study their ways of evolution and other relevant aspects, as open access, friendly interfaces, accessibility, standardization level, interdisciplinary work and interoperability. The final aim of these analysis and reflections is to offer new guidelines and perspectives for the future of the text, taking it further than a simple study object.

Keywords: Digital scholarly edition. Literary texts. Interoperability. Standard. Encoding system.

Palabras clave: Edición digital académica. Textos literarios. Interoperabilidad. Estándar. Codificación.

EL PUNTO DE PARTIDA

La edición digital de textos es un tema que se encuentra en el punto de mira por su situación cambiante, por las amplias perspectivas que ofrece gracias a la tecnología y por los diversos sectores que se ven implicados en el proceso. Los retos de la edición digital han sido tema de reflexión de importantes trabajos (ver, por ejemplo el magnífico artículo de Boot o el trabajo de Buzzetti y McGann, o, en el ámbito nacional, Fradejas). Si bien nos encontramos en un entorno en el que tecnológicamente las posibilidades son casi infinitas, lo cierto es que los datos que la realidad nos ofrece distan aún mucho de ese ideal posible al que debemos aspirar.

Hablar de la tradición editorial académica en España podría dar lugar a varios trabajos enciclopédicos. Nuestra tradición, arraigada desde las prensas de Hurus en los primeros períodos de producción incunable hasta la actualidad, ha sido muy larga y productiva. La historia cuenta con figuras de grandes editores desde Pascual de Gayangos y Menéndez Pidal, a figuras actuales que destacan por su cuidado en la edición y por sus valiosos resultados, como Pedro Sánchez Prieto-Borja,¹ Miguel Ángel Pérez Priego o Pedro Cátedra, entre muchos otros, además de muchos grandes proyectos de investigación dedicados a la edición de textos literarios.²

Este trabajo no va a ocuparse, sin embargo, de esa edición de corte tradicional, sino de las nuevas perspectivas de futuro en la sociedad actual, en la sociedad del Conocimiento y de la Información de datos interconectados cuya presencia en el mundo digital nos invade y transforma nuestros modos de vida y pensamiento.³ Esta transformación ha resultado evidente y necesaria en ciertos sectores, como la industria y la medicina, que han visto en la tecnología el aliado clave para el avance en la investigación y en su desarrollo.⁴ Los frutos de la alianza tecnológica con otras áreas de investigación se han dejado sentir notablemente y no van a ser objeto aquí de comentario. Las disciplinas

-
1. Investigador, sin embargo, muy concienciado por la importancia del mundo digital en la filología tradicional, como demuestran sus publicaciones, además de la iniciativa Red Charta, que él coordina.
 2. Por poner un ejemplo, el proyecto TC-12 Consolider sobre Patrimonio Teatral Clásico Español, desarrollado en los últimos años, que ha dado lugar a numerosas ediciones de textos clásicos del drama del Siglo de Oro.
 3. En este sentido, resulta altamente ilustrativo el libro de Manovich, muy documentado, que define el software como una cola de pegar que está en el ambiente e impregna todo.
 4. Entre otras referencias, cabe destacar la obra de Negroponte, que a pesar de los años que lleva escrita, continúa siendo de rabiosa actualidad.

humanísticas, sin embargo, han quedado rezagadas en esta suerte evolutiva por varias razones: en primer lugar, por la propia naturaleza de su objeto de estudio, cuya utilidad inmediata es la difusión cultural y el aprendizaje, pero sin ofrecer una rentabilidad económica en su explotación directa, y en segundo lugar, por el perfil de los investigadores y de los sistemas de investigación dedicados al estudio de las humanidades, caracterizados por su escasa interdisciplinariedad, por su método de trabajo principalmente autónomo, por contar con un presupuesto por lo general bajo y por tener poco presente en sus métodos de análisis el componente tecnológico.

Recientemente han comenzado a despertar las voces que claman la necesidad de adecuación de las Humanidades a esta tendencia global. Por citar algunos datos que muestran esta apertura hacia nuevos enfoques, destacan, entre otros, el informe de la *American Academy for Arts and Sciences* o el proyecto italiano *New Humanities*, cuyas características y aportaciones detallo en otro lugar (González-Blanco 53).

Resulta además significativo el aumento de cifras de publicaciones en formato electrónico y especialmente de contenidos en abierto a través de la web, que, junto a proyectos de digitalización masiva como Google Books, han provocado que el mercado editorial en el libro impreso descienda considerablemente en comparación con el formato electrónico⁵ y haya hecho a este sector empresarial replantearse nuevas vías de supervivencia explotando otras opciones, como el desarrollo de los contenidos multimedia, de portales educativos⁶ o de *e-books* (Zanoni).

Otro de los factores que ha influido notablemente en la modificación de los planteamientos ha sido la iniciativa de las bibliotecas. La falta de espacio físico unida a las necesidades de digitalización y difusión de contenidos, el aumento de recursos electrónicos y bases de datos orientadas a la investigación, y el llamamiento internacional a la interoperabilidad⁷ y a la difusión de los contenidos en abierto han convertido a estas instituciones en uno de los promotores de la adopción de nuevos modelos de contenido electrónico.

5. 2011 es el primer año en el que el volumen de libros electrónicos supera al de impresos.

6. Como ejemplo tenemos a la editorial SM, que innova constantemente para generar distintos recursos y plataformas educativas, como <http://www.librosvivos.net/portada.asp>.

7. Las normas BS 8723 y ANSI/NISO Z39.19 definen la interoperabilidad como la capacidad de dos o más sistemas o componentes de intercambiar información y de usar esa información que se ha intercambiado.

LA EDICIÓN DIGITAL EN ESPAÑA

Si bien en los sectores descritos en nuestro país nos encontramos en la línea de desarrollo marcado por las necesidades de internacionalización y los intereses económicos y comerciales, la situación es muy diversa en determinados ámbitos, como el de la edición de textos literarios, que, en comparación con lo que sucede en otros países resulta pobre y cuenta con proyectos independientes, realizados en códigos y con tecnologías diferentes que hacen que no sean interoperables entre ellos, formando un mosaico de formas diversas que hacen que, en muchos casos, los propios investigadores desconozcan su existencia.

El campo de la edición de textos literarios castellanos es una muestra significativa de la denominada “brecha digital”, que separa a los grandes editores de la filología tradicional, que continúan haciendo magníficas ediciones en papel, contrastando con el mosaico de ediciones presente en la nube, muy poco significativo en volumen y mal valorado aún desde el ámbito académico (Morrás).

Esta visión más bien pesimista de lo que queda por hacer contrasta, sin embargo, con algunos proyectos pioneros cuya mención no se puede obviar, por haber sido y ser grandes puntos de referencia que han ido haciendo cambiar poco a poco nuestras formas de investigación. En lo que respecta a los grandes proyectos de edición digital, en España contamos con:

–El CORDE,⁸ de la Real Academia Española, pionero por su temprana fecha de inicio y por su sistema de etiquetado utilizando la Text Encoding Initiative TEI (con una DTD basada en SGML, tal y como señala Martín de Santa Olalla), es el corpus de textos antiguos de la Real Academia, utilizado especialmente para consultas de tipo lingüístico. Se trata de una herramienta que no realiza ediciones digitales propiamente dichas de los textos, sino que facilita la búsqueda de datos contextualizados.

–La *Biblioteca Virtual Cervantes*: es el repositorio más rico de textos del que disponemos, ampliamente utilizado para fines didácticos, de investigación y de difusión. Si bien se elaboró siguiendo los estándares de edición que marcaba el consorcio TEI, hay que decir, sin embargo, que el acceso al etiquetado de estos textos más allá de la mera lectura

8. Ver las direcciones electrónicas en el listado anejo.

no está disponible en abierto para su reutilización por el lector o el investigador. Presenta además el problema de que los contenidos de conjunto son muy amplios pero también muy variados. Esta biblioteca está desarrollando actualmente un sistema de “linked data” y una ontología que le permitirá en breve, dar un paso más allá, tal y como muestra el premio que recientemente se les ha otorgado.⁹

Además de estos dos grandes repositorios digitales de textos existen algunos proyectos menos extensos en cuanto a alcance o temática (segmentados por época, géneros o finalidad), como son *Parnaseo*, que incluye varias bases de datos además de ediciones críticas digitales y ediciones de textos en formato pdf y facsimilar; *Aul@medieval*, dentro de Parnaseo, que también incluye ediciones digitales; *Tesoro*, proyecto de la Universidad Carlos III de Madrid que digitaliza textos de teatro del Siglo de Oro utilizando tecnologías XML, o *Arte-lope*, que incluye una base de datos de textos del Siglo de Oro y una colección de ediciones digitales en XML. Es interesante también citar *e-Spania*, revista digital creada en 2006 que publica ediciones de textos basada en la plataforma *OpenEdition*.

Estos proyectos presentan el mismo problema: utilizan herramientas personalizadas para sus propios prototipos pero no comparten su código de forma pública ni sus esquemas de codificación, por lo que se reinventa la rueda en cada uno de los proyectos, de forma que la dependencia del soporte informático es total y la interoperabilidad y reutilización de sistemas de codificación entre unos y otros, baja.

LAS GRANDES PLATAFORMAS Y PROYECTOS INTERNACIONALES DEDICADOS A LA EDICIÓN

La situación internacional amplía bastante las perspectivas en lo que se refiere a la oferta de proyectos y herramientas. Existe una gran variedad de recursos y plataformas y la mayoría de proyectos grandes se esfuerzan por seguir el camino hacia los estándares con el fin de compartir información y código para hacerlos interoperables (interoperabilidad que, por otro lado, aún están lejos de alcanzar en su mayoría, ya que la variedad de estándares disponibles

9. Se trata de un prestigioso premio para innovación en el ámbito bibliotecario concedido por la universidad de Stanford, cuya información puede verse en: <http://library.stanford.edu/projects/stanford-prize-innovation-research-libraries-spirl/2013-spirl-winners>.

es enorme y no hay acuerdo común incluso dentro de los mismos estándares, tal y como señala Bauman). Son frecuentes los proyectos que depositan su código en *GitHub*, u ofrecen descripciones detalladas de sus sistemas de codificación a través de la web. El principal problema que se plantea en este punto es la dificultad por parte de los investigadores de conocer y disponer de toda esta información de forma consolidada, ya que no existe un repositorio informativo actualizado que contenga el conjunto de ediciones y herramientas que se están desarrollando. Ha habido en esta línea varias iniciativas, como el proyecto *Bamboo* (ver Dombrowski/Denbo), al parecer hoy fracasado, el portal Arts&Humanities.net en el ámbito anglosajón (recientemente vinculado a Centernet) o el proyecto *Tapas-TEI* de reciente creación orientado a crear un repositorio de código TEI (ver Flanders/Hamlin).

En lengua española carecemos totalmente de una perspectiva informativa en esta línea, por lo que el presente trabajo va a tratar de reunir y describir algunos de los proyectos más significativos que pueden resultar de interés para el avance de la edición digital de textos en nuestro país.

Comenzaré por citar algunas de las plataformas digitales que además de la propia colección de textos editados contienen herramientas u otro tipo de recursos para su explotación y utilización:

En primer lugar, en Alemania destaca *TextGrid*. Es un proyecto que aúna diez instituciones o grupos de trabajo, cuyo objetivo es establecer una infraestructura adecuada para un entorno de investigación virtual en humanidades. Ofrece herramientas y servicios de acceso libre y gratuito y dispone de un Laboratorio en el que se trabaja y se ofrecen estas herramientas, un repositorio, para almacenar de manera segura, para publicar o para investigar textos, imágenes, bases de datos, y otro tipo de información.

Las iniciativas editoriales en Alemania se extienden al ámbito de las academias, entre las que destaca la Berlin-Brandenburg Academy of Sciences and Humanities (BBAW), que se ha dotado de un potente grupo de digitalización que ha creado diferentes recursos relacionados con la edición académica digital. Cabe destacar la iniciativa *Ediarum*, un conjunto de tecnologías que proporcionan un entorno virtual de trabajo en el cual los manuscritos y otros documentos marcados en XML-TEI pueden ser transcritos y editados. Los inconvenientes del paquete que ofrecen comienzan por su difícil instalación, a la que se añade que solo se ofrece en alemán y el editor debe conocer múltiples lenguajes informáticos relacionados con su funcionamiento, lo que convierte esta excelente herramienta en elitista por estar al alcance de unos pocos.

En Holanda se han realizado en esta línea numerosos proyectos, entre los que destaca la plataforma *eLaborate*, que ofrece herramientas para edición de textos en TEI-XML con distintas posibilidades de visualización y búsqueda de contenidos. En este momento están trabajando sobre la tercera versión de esta plataforma, pero aún no se conocen muchos detalles sobre sus características.

En la Universidad de Birmingham se ha desarrollado el proyecto de *Scholarly Digital Editions* liderado por Peter Robinson, que utiliza las herramientas digitales para realizar el proceso de “collatio” textual, como *Collate* y *Anastasia electronic publishing system* (consistente en un sistema analítico de herramientas de integración SGML/XML).

En Italia destaca la plataforma *Murruca* usada para crear bibliotecas digitales. Se trata de un marco editorial que permite la edición digital y la gestión de colecciones interoperables con otros proyectos gracias a estar basada en un modelo semántico de Linked Data. Es simple, modular y de fácil manejo. Permite editar y enlazar material digital, pero es también una herramienta que puede emplearse en la investigación. Admite fuentes primarias y secundarias (manuscritos, cartas, imágenes, audio, fotografías, vídeos) en distintos formatos (xml, doc, odt, pdf, tiff, gif, jpeg, png, mpeg, avi, flv). Contiene además otras herramientas como *Pundit* y *Flexip* cuya función es dotar al texto de contenidos semánticos.

En este marco de grandes proyectos y plataformas ha surgido recientemente una Acción COST denominada *Interedition*, cuyo objetivo es facilitar la interoperabilidad de herramientas y metodologías empleadas en el ámbito de la edición académica digital de textos y en la investigación en dicho terreno. Desde 2014 cuentan con un proyecto llamado *Microservices*,¹⁰ consistente en ofrecer pequeños servicios web en la nube que desarrollan tareas específicas del ámbito académico. Su objetivo es promover la interoperabilidad y sostenibilidad de herramientas digitales académicas con las que trabaja el grupo de *Interedition*.

En Francia, la plataforma que se ha convertido en una referencia internacional es *Openedition*; es una biblioteca digital de grandes dimensiones que alberga en su interior cuatro portales: *Openedition Books*, que publica colecciones de libros, *Revue.org*, que se dedica a la edición digital de revistas, *Calenda*, que recoge actividades y eventos académicos, e *Hypothèses*, plataforma

10. http://www.interedition.eu/wiki/index.php/About_microservices.

de blogging científico. Los textos que recoge están editados en XML con un software denominado *Lodel*. Esta plataforma ha sido sin duda uno de los grandes pasos hacia la edición y difusión de los textos digitales en abierto.¹¹

LOS PROYECTOS INNOVADORES DE EDICIÓN DIGITAL EN EL ÁMBITO INTERNACIONAL

Dentro de los portales específicos dedicados a la edición de textos destaca en Holanda el creado por el Centrum voor Teksteditie en Bronnenstudie¹² denominado *Van Nu en Straks. De Brieven*, que ofrece disponible toda la información de metadatos de TEI a modo de una barra lateral de etiquetas, el código XML-TEI, la imagen digitalizada del texto que sirve de punto de partida a la edición y diferentes tipos de edición del texto (diplomática y modernizada), además de una información sobre los personajes mencionados al pasar el cursor por encima del texto.

También son interesantes en el ámbito holandés las ediciones realizadas a partir de la herramienta ya citada *e-Laborate*, como esta edición de *Bartholomeus Engels*: <http://bartholomeusengelsman.huygens.knaw.nl/path>, con distintas posibilidades de muestra de anotaciones, traducción de palabras, referencias lingüísticas y sistemas de visualización de imágenes de los manuscritos.

Otro proyecto interesante en esta línea es el también holandés *Samuel Beckett Digital Manuscript Project*, financiado por el ERC, más centrado en edición genética y con herramientas de visualización por coordenadas que permiten situar la transcripción del propio texto sobre la imagen y además ofrecen información muy rica sobre el proceso compositivo de la realización de la edición.

En una línea más colaborativa y abierta a la edición grupal quiero destacar en Reino Unido el proyecto *Transcribe Bentham*, del University College of London, que han llegado a utilizar el *crowdsourcing*¹³ como sistema colaborativo para la edición digital. En este proyecto resulta muy interesante el

11. Filosofía que además sigue las pautas del Consejo de Europa y EBLIDA sobre legislación y política bibliotecaria en Europa, que pueden consultarse en <http://www.eblida.org/>.

12. Además, destacan las distintas herramientas de apoyo a la edición digital ofrecidas por este centro que pueden consultarse en <http://ctb.kantl.be/>.

13. Este término hace referencia a la colaboración colectiva mediante contribuciones participativas en la edición del texto. Este es el modelo que sigue, por ejemplo, la Wikipedia.

que, a pesar de haber etiquetado el texto con TEI, han tratado de simplificar al máximo el proceso de edición para hacerla lo más colaborativa posible, cambiando las etiquetas por botones de iconos, de forma que para un usuario no familiarizado con sistemas de etiquetado XML es muy sencillo trabajar de forma casi intuitiva y visual.

También de Londres, del Department of Digital Humanities del King's College parte la iniciativa de la plataforma *Out of the Wings*, que aunque no es propiamente una plataforma dedicada a la edición digital, alberga proyectos como la edición digital de *La Entretenida*, que supone un avance importante en la edición digital de textos literarios en castellano: <http://entretenida.ou-tofthewings.org/index.html> (ver Spence).

La idea de ofrecer un sistema de interfaz amigable al usuario se repite en el proyecto *T-Pen*, de la Universidad de St. Louis (EE.UU.): *Transcription for editorial and paleographical notation*. Está orientado a la transcripción paleográfica de manuscritos, también con una presentación de etiquetas de una forma sencilla y adaptada a todos los niveles en un entorno virtual colaborativo. El código sin embargo, no se encuentra disponible en abierto.

En este sentido de simplificación de la interfaz es también notable el esfuerzo realizado por el gran proyecto alemán *Deutsches Textarchiv*, que etiqueta sus textos en TEI pero incluye una herramienta de personalización del software de edición *Oxygen* con el fin de facilitar a los usuarios la inteligibilidad de las etiquetas en lenguaje humano y mediante iconos gráficos.

LAS HERRAMIENTAS

Existen numerosas herramientas para el tratamiento de textos en XML, las más utilizadas se encuentran detalladas en la plataforma de *Interedition* ya citada.¹⁴ De todas ellas, quiero destacar y hacer una breve reseña de las que resultan más significativas por sus posibles aplicaciones e implementaciones en plataformas de edición digital.

En el marco del estándar de la *Text Encoding Initiative* TEI es conocida la herramienta de TEI *Boilerplate*, que transforma de manera muy simple los documentos XML a través de XSLT mediante el uso de unas plantillas preins-

14. http://www.interedition.eu/wiki/index.php/Existing_Tools. A esta lista se suma un post de discusión en Academia.edu sobre las herramientas mejores para edición digital: https://www.academia.edu/Questions/8491/What_good_critical_edition_and_collation_software_is_the_re_for_medievalists_Thank_you.

taladas. Este recurso es especialmente útil para humanistas que conocen el sistema de etiquetado en XML pero no son capaces de autogenerar hojas de transformación XSLT.

El ámbito de la “collatio” textual o la comparativa automatizada entre textos con elementos comunes ha sido desde los inicios una de las áreas que más han llamado la atención al campo de la informática. En esta línea destacan proyectos como *Collatex*, desarrollado por el *Interedition Development Group*, específicamente diseñado para la “collatio” de los textos en crítica genética y la elaboración del *stemma* y el aparato crítico de una forma semiautomática basada en la diferencia de variantes. La herramienta es de código abierto (disponible en *GitHub*) y trabaja basada en TEI y en representación de grafos de las relaciones entre las variantes del texto, una vez sometido este a un proceso de tokenización.

Collate es un programa para comparar y transcribir versiones de los textos capaz de llevar a cabo la *collatio* de varias copias manuscritas (transcritas) de un texto. Presenta una interfaz orientada a objetos, para hacer el trabajo más cómodo al autor y por la posibilidad de obtener con ello una configuración completa.

En la misma línea pero de carácter propietario y como aplicación de escritorio, se ha desarrollado el software del *Classical Text Editor*, software muy bien desarrollado, basado en TEI, aunque con una metodología de trabajo muy diferente, pues orienta más sus funciones hacia las de editor que aspira a obtener un producto final en un formato más orientado hacia el libro tradicional. Como herramienta es muy completa, pero como planteamiento y aplicación está quedando un poco desfasada dentro de un panorama global en el que el acceso libre, el código abierto y la interoperabilidad son las máximas que rigen la mayor parte de los proyectos.

De menor alcance, pero también de importancia citaré *Juxta*, software que sirve también para editar textos y comparar versiones, sin embargo, sus funcionalidades para la edición académica compleja son limitadas.

En la línea de los editores fuera del ámbito anglosajón destaca el proyecto *Edictor*, desarrollado por la Universidad de São Paulo, que ofrece interesantísimas perspectivas sobre todo por su propia génesis dentro de un ámbito cultural distinto en el que las Humanidades Digitales están en auge y desarrollo.

Sin ser un proyecto de edición digital propiamente dicho, *Versioning Machine 4.0*, desarrollado por Susan Schreibman, permite comparar distintas ver-

siones de una edición a lo largo del tiempo de forma que el editor pueda ser consciente de las consecuencias de su toma de decisiones. Esta herramienta adolece de los mismos problemas que la anterior: resulta insuficiente para un tipo de proyecto filológicamente complejo, en el que las técnicas han de afinarse, sin embargo, cabe destacar que la idea es excelente y de gran utilidad sobre todo para ediciones de tipo genético en las que el proceso editorial y la evolución del propio autor resultan fundamentales.

En la línea de software más técnica –y que probablemente sorprenda a los lectores– se encuentra *Latex*, software muy conocido y específicamente diseñado para las ciencias, pero que ofrece interesantísimas posibilidades para la edición (pues incluye un programa de composición de textos).¹⁵

También software, pero con un perfil muy diferente, tenemos *Scripto*, una herramienta libre y de código abierto que facilita la transcripción colaborativa de documentos y archivos multimedia. Su distinto perfil se debe a que fue diseñada para que instituciones y organizaciones como bibliotecas y museos participen en proyectos a gran y pequeña escala, y tanto en pequeños grupos como en proyectos individuales. Una de sus ventajas es que es compatible e integrable con los principales CMS de código abierto, como Omeka, Wordpress y Drupal.

En la línea de las herramientas colaborativas está también *Textual Communities*, plataforma desarrollada en la Universidad de Saskatchewan (Canadá), concebida como un espacio de trabajo en equipo donde los colaboradores puedan realizar transcripciones en un entorno virtual donde dispongan de la imagen que les sirve de fuente. Esta filosofía del *crowdsourcing* se ha llevado a cabo en muchos proyectos de edición textual, sobre todo en el ámbito americano. Por citar uno de ellos, es especialmente completo el archivo digital de Martha Berry, en el que además se dan “premios” virtuales a los colaboradores participantes, cuya experiencia se gradúa a través de distintos niveles.

La anotación de textos, manuscritos y otras fuentes no textuales como imágenes u objetos artísticos es la base de un gran proyecto denominado *Open Annotation*, que está desarrollando un modelo de ontología colaborativa para

15. Sobre la edición digital con Latex, son muy interesantes las recomendaciones de Uwe Lück en http://www.webdesign-bu.de/uwe_lueck/criteditlx.html. En Italia se realizan numerosas ediciones con este software, como puede verse, como ejemplo en la actividad desarrollada por el grupo GUIT, o en los diversos cursos de edición de texto con esta herramienta organizados por el profesor Maurizio Lana: <https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=digitalclassicist;d3057e9.1312>.

permitir la interoperabilidad entre textos e imágenes anotadas mediante metadatos y aspira a convertirse en un estándar dentro del W3C. Una aplicación concreta de este tipo de anotación puede verse en el proyecto *Shared Canvas*, herramienta para anotar colaborativamente mediante metadatos el mismo objeto digital con la posibilidad de exportarlo, explotarlo y enlazarlo con otros recursos. El potencial de estos proyectos es muy grande, ya que amplía las posibilidades de localizar, por ejemplo, piezas del mismo manuscrito localizadas en diferentes bibliotecas, principio en el que se están basando ya algunos proyectos, como *Perseids* para los fragmentos clásicos.

Aunque su finalidad dista más de ser una herramienta de edición de textos, quiero citar aquí para finalizar *Philologic*, desarrollado por ARTFL Project en la Universidad de Chicago, en colaboración con el Centro de Desarrollo Digital de la Biblioteca de Chicago. Es una herramienta complementaria a las de edición que permite la búsqueda en una amplia variedad de bases de datos de la web. Es fácil de usar, pero potente, realiza búsquedas a texto completo, recuperación de textos, e incluye un sistema para mostrar los contenidos de bases de datos multimedia (textos, imágenes, audio) con la posibilidad de analizar estructura de textos complejas (ej. SGML, BetaCode), y con gran número de metadatos. Sus funcionalidades están diseñadas para la investigación académica sobre bases de datos de textos literarios, religiosos, filosóficos e históricos.

LOS NUEVOS RETOS DE LA EDICIÓN DIGITAL: PROPUESTAS Y SOLUCIONES

La anterior descripción de herramientas y recursos permite extraer varias conclusiones que ya se han ido apuntando a lo largo del trabajo pero que sirven para iluminar nuevas vías de conocimiento científico de cara al avance de la investigación en el mundo académico editorial. Realizando un análisis evaluativo y comparado entre los diferentes recursos descritos, se observa que las tendencias principales en la edición digital global apuntan a una serie de principios y realidades que todo nuevo editor que comienza un proyecto digital debería conocer y quizás practicar:

–En primer lugar, es necesario realizar un trabajo “de campo” que en este caso sería más propiamente “de navegación” en el que merece la pena invertir tiempo para analizar con detalle las diferentes plataformas, herramientas y proyectos como los que aquí se han descrito que

puedan servir como modelos y como ideas inspiradoras de lo que el editor desea hacer.

- Es fundamental asumir, como punto de partida, que la tecnología aplicada a la edición de textos se encuentra enormemente desarrollada, por lo que es muy posible que lo que el editor necesite sea una adaptación o una implementación de sus necesidades sobre un recurso ya existente, lo que permitirá, sin duda, el enriquecimiento de la herramienta y el avance del conocimiento. Toda propuesta de empezar de cero supone una pérdida de tiempo, aunque también implica un aprendizaje.
- Los estándares existen y deben conocerse y aplicarse, con el fin de fomentar la interoperabilidad entre distintos proyectos y la preservación del contenido digital. En el campo de la edición de textos es fundamental el conocimiento del modelo TEI, desarrollado por el Consorcio TEI desde 1987 y en constante desarrollo y mejora.
- Para elaborar una edición, herramienta o recurso digital, es necesario, en primer lugar, preguntarse por su finalidad, tener muy claro dónde se quiere llegar, e incluso elaborar un prototipo del proyecto ideal que el editor tiene en mente, materializándolo y compartiéndolo con los miembros de su equipo. La improvisación y las modificaciones en terrenos digitales son costosas en tiempo, dinero y trabajo, y aunque siempre son necesarias, minimizarlas debe ser el objetivo.
- La edición digital de textos se difunde hoy día en internet. Realizar ediciones en otros soportes como el CD o aplicaciones de escritorio reduce enormemente las posibilidades de interconexión entre proyectos. La web es, por lo tanto, el lugar ideal de difusión de contenidos y datos.
- Se debe aspirar a escribir en un lenguaje entendible e interpretable por máquina que sea fácilmente codificable y traducible al HTML de la web, razón por la cual se imponen cada vez más los sistemas basados en codificación XML (como TEI para el etiquetado, XSLT para la transformación o RDF para la definición de metadatos y su interrelación como datos enlazados o *Linked Data*). Estos formatos tienen además la ventaja de su

fácil exportación, transformación y preservación gracias al poco espacio que ocupan al ser almacenados en los servidores como texto.

- Es necesario destronar el PDF como formato idóneo para la edición digital (Hipola/Brun), pues no es sino una emulación de la hoja de papel a través de la pantalla y, a pesar de su extendido uso, presenta posibilidades muy limitadas en lo que a la explotación del texto se refiere, que se reducen a la realización de búsquedas concretas. La utilización de lenguajes basados en XML –eXtensible Markup Language– permiten, como su nombre indica, ampliarlos hasta niveles de gran profundidad que ofrecen amplísimas posibilidades de explotación y transformación del texto, amén de ser el formato de intercambio de datos en la red por excelencia y de permitir el reflejo de estructuras semánticas internas dentro de los documentos. Además, este tipo de codificación añade su sencilla transformación a soportes multiformato, como los utilizados para los ebooks (ePub, por poner un ejemplo). Una ventaja añadida es que la accesibilidad es mucho mayor a través de estos formatos, ya que se pueden incorporar de forma sencilla imágenes, archivos de audio u otros recursos en caso de necesidad y en cualquier momento o lugar.
- La difusión en abierto del producto o resultado es la forma ideal de publicar las ediciones digitales en la web. Este es probablemente el punto más controvertido, dados los intereses editoriales y el modelo económico de cambio que arrastra muchas implicaciones en este sentido en lo que a los datos en abierto respecta, pero la realidad hoy día es que los textos que más se leen son los que se encuentran en abierto en internet. La plataforma *Openedition* ya citada es un gran modelo a seguir. Por otro lado, es importante definir y enmarcar las ediciones con licencias, como las *Creative Commons*, específicamente diseñadas para proteger los derechos de autor en publicaciones, que permiten a los autores definir legalmente y sin coste el grado de explotación de su obra, que abarcan desde el *non commercial* hasta el *copyleft* para la libre explotación de los textos. El problema que estas licencias plantean es que son, en muchos casos poco conocidas y por ello, mal utilizadas.
- Es necesario asumir que los nuevos procesos de edición digital implican nuevas formas de trabajo con dos características importantes:

el trabajo ha de realizarse en equipos interdisciplinares en los que se establezca el continuo diálogo entre editores y técnicos, y la colaboración y participación externa resulta fundamental a la hora de explotar los recursos y las herramientas creadas para poder obtener resultados satisfactorios. El modelo del *crowdsourcing*, bien conocido por el efecto *Wikipedia*, es un modelo a trasladar e incorporar a nuevos campos, como ya algunos proyectos han hecho.¹⁶

–Desde el punto de vista tecnológico también el potencial de mejora es amplio, pues se han de desarrollar sistemas de modelado del conocimiento, así como tratamientos semánticos de la información aplicados a campos novedosos en el sector, y en especial se ha de trabajar en la accesibilidad y el diseño de interfaces con el fin de crear un entorno digital amigable a los humanistas sin necesidad de asumir gran cantidad de información tecnológica. Se debe aspirar a reducir la brecha digital de dos formas: en primer lugar, fomentando la creación de interfaces amigables y sencillas que faciliten la explotación de los distintos niveles del texto, y en segundo lugar, ofreciendo información y formación sobre posibilidades y perspectivas tecnológicas a los humanistas encargados de la edición al nivel textual.

–Uno de los objetivos del editor debe ser la total transparencia del proceso de elaboración, realizando de forma paralela una “hoja de ruta” de todo el proceso editorial, que sirva de guía tanto a los participantes del proyecto como a los lectores. Además, se recomienda poner el código utilizado a disposición de la comunidad científica, disponible a través de la propia página web o bien en repositorios específicamente habilitados para ello, como *GitHub*. La disponibilidad del código fuente posibilita la continuación, mejora y avance del conocimiento en otros proyectos además de la interoperabilidad y la posible integración de los mismos en buscadores semánticos y en la red, cada vez más tendente a seguir los principios de *Linked Open Data* (ver Vandenbussche/Vatant).

16. En este sentido, resulta muy ilustrativo el libro de Cobo Romaní y Pardo Kuklinski, que estudia las perspectivas del fenómeno de colaboración y el alcance de la web 2.0 o web social.

–La edición digital entendida en esta dirección va mucho más allá de visualizar el propio texto como producto final de la cadena editorial. El propio proceso de lectura está condicionado por las distintas posibilidades que las herramientas ofrecen al lector, que podrá tener, en muchos casos a su disposición diferentes opciones de visualización del contenido, de selección de variantes, de notas aclaratorias de tipo lingüístico o explicativo, e incluso la posibilidad de disponer del manuscrito o impreso original junto al texto a modo de imagen digitalizada. El lector se convierte, de esta forma, en el “dueño” del texto, que si bien conserva intacta su estructura, puede ser anotado, convertido, visualizado, adaptado a nuevos formatos, trasladado y luego recuperado según las necesidades del lector. Esta idea además tiene unas amplias posibilidades de desarrollo futuro para el fomento de la lectura dentro de una sociedad como la actual, en la que la inmediatez de la visualización y el poder de la imagen van en detrimento del conocimiento lineal aportado por el renglón que ha estructurado nuestras cabezas y las de nuestros padres pero no las de los nuevos nativos digitales (ver Brea).

–Las aplicaciones de este tipo de tecnologías van mucho más allá y presentan unas perspectivas amplísimas a partir de los textos editados y etiquetados de esta forma donde aún hoy hay mucho trabajo para humanistas, libreros, filólogos y bibliotecarios y numerosas aplicaciones útiles para instituciones como academias, bibliotecas, archivos y museos. Ejemplo de ello es la creación de aplicaciones para el turismo, la combinación de sistemas de estadística, visualización y representación geográfica y el estudio conjunto e interdisciplinar de resultados que, en la sociedad de la información actual en que vivimos, regida por los enlaces de grandes volúmenes de datos, aún queda mucho trabajo por hacer.¹⁷

RECURSOS DIGITALES CITADOS

Anastasia electronic publishing system <<http://www.sd-editions.com/anastasia/index.html>>

17. Quiero dar las gracias a Susanna Allés, Elena Azofra, José Manuel Fradejas, Gimena del Río, Clara Martínez, María Martos y Jesús Urueña por sus sugerencias y aportaciones en la recogida de material que ha servido de base a la elaboración de este artículo.

Archivo digital de Martha Berry <<https://mbda.berry.edu/>>
 Artelope <<http://artelope.uv.es/>>
 Arts&Humanities.net <<http://www.arts-humanities.net/>>
 Aul@ Medieval <<http://parnaseo.uv.es/@Medieval.html>>
 Bamboo Project <<http://www.projectbamboo.org/>>
 Biblioteca Virtual Miguel de Cervantes <<http://www.cervantesvirtual.com/>>
 Centrum voor Teksteditie en Bronnenstudie <<http://ctb.kantl.be/>>
 Classical Text Editor <<http://cte.oew.ac.at/>>
 Collate <<https://metacpan.org/pod/Text::TEI::Collate>>
 Collatex <<http://collatex.net/>>
 Corde <<http://corpus.rae.es/cordenet.html>>
 COST Action Interedition <<http://www.interedition.eu/>>
 Creative Commons <<http://es.creativecommons.org/blog/>>
 Deutsches Textarchiv <<http://www.deutschestextarchiv.de/>>
 e-Spania <<http://e-spanialivres.revues.org/?lang=es>>
 EBLIDA <<http://www.eblida.org/>>
 Ediarum <<http://www.bbaw.de/telota/software/ediarum>>
 Edictor <<http://humanidadesdigitais.org/edictor/>>
 eLaborate: <<http://www.e-laborate.nl/en/>>
 GitHub <<https://github.com/h>>
 GUIT <<http://www.guit.sssup.it/latex/critical.php>>
 Juxta <<http://wiki.tei-c.org/index.php/JuxtaWS>>
 LaTeX <<http://www.latex-project.org/>>
 Lodel <<http://lodel.org/>>
 Muruca <<http://www.muruca.org/>>
 Open Annotation <<http://www.openannotation.org/>>
 Open Edition <<http://www.openedition.org/>>
 Out of the Wings <<http://www.outofthewings.org/>>
 Oxygen <<http://www.oxygenxml.com/>>
 Parnaseo <<http://parnaseo.uv.es/>>
 Perseids <<http://sites.tufts.edu/perseids/>>
 Philologic <<https://sites.google.com/site/philologic3/home>>
 Red Charta <www.charta.es>
 Samuel Beckett Digital Manuscript Project <<http://www.beckettarchive.org/>>
 Scholarly Digital Editions <<http://www.sd-editions.com/>>

Scripto <<http://scripto.org/>>
 Shared Canvas <<http://www.shared-canvas.org/>>
 SM Editorial <<http://www.librosvivos.net/portada.asp>>
 T-Pen <<http://t-pen.org/TPEN/>>
 TAPAS-TEI <<http://www.tapasproject.org/>>
 TC-12 Consolider <<http://tc12.uv.es/>>
 TEI Boilerplate <<http://dcl.slis.indiana.edu/teibp/index.html>>
 Tesoro <<http://www.bib.uc3m.es/~nogales/xml/tesoro/>>
 Text Encoding Initiative TEI <<http://www.tei-c.org/>>
 TextGrid en Alemania <<http://www.textgrid.de/>>
 Textual Communities <<http://www.textualcommunities.usask.ca/es/web/textual-community>>
 Transcribe Bentham <<http://www.transcribe-bentham.da.ulcc.ac.uk>>
 Van Nu en Straks. De Brieven <<http://www.vnsbrieven.org/>>
 Versioning Machine 4.0 <<http://v-machine.org/>>

SIGLAS

ARTFL: Analyse et Traitement Informatique de la Langue Française
 BBAW: Berlin-Brandenburg Academy of Sciences and Humanities
 CMS: Content Management System
 CORDE: Corpus Diacrónico del español
 EBLIDA: European Bureau of Library, Information and Documentation Associations
 ERC: European Research Council
 RDF: Resource Description Framework
 ReMetCa: Repertorio Métrico Digital de la Poesía Castellana Medieval
 TAPAS: TEI Archiving Publishing and Access Service
 TEI: Text Encoding Initiative
 SGML: Standard Generalized Markup Language
 XML: Extensible Markup Language
 XSLT: Extensible Stylesheet Language Transformations

OBRAS CITADAS

Bauman, Syd. "Interchange vs. Interoperability". *Proceedings of Balisage: The Markup Conference 2011. Balisage Series on Markup Technologies* 7 (2011).

- doi:10.4242/BalisageVol7.Bauman01. 17 de abril de 2014. <<https://balisage.net/Proceedings/vol7/html/Bauman01/BalisageVol7-Bauman01.html>>
- Brea, José Luis. *Cultura RAM: mutaciones de la cultura en la era de su distribución electrónica*. Barcelona: Gedisa, 2007.
- Boot, Peter. "Some Digital Editions and Some Remaining Challenges". *Janus Digital* (2012): 39-54. 31 de enero de 2014. <<http://www.janusdigital.es/articulo.htm?id=7>>.
- Buzzetti, Dino, y Jerome McGann. "Critical Editing in a Digital Horizon". *Electronic Textual Editing*. Eds. John Unsworth, Katherine O'Brien O'Keeffe, Lou Burnard. New York: MLA, 2006. 37-50.
- Cobo Román, Cristóbal, y Hugo Pardo Kuklinski. *Planeta Web 2.0. Inteligencia colectiva o medios fast food*. Barcelona/México DF: Flacso, 2007.
- Dombrowski, Quinn, y Seth Denbo. "TEI and Project Bamboo". *Journal of the Text Encoding Initiative* 5 (junio 2013). 31 de enero de 2014. <<http://jtei.revues.org/787>>.
- Flanders, Julia, y Scott Hamlin. "TAPAS: Building a TEI Publishing and Repository Service". *Journal of the Text Encoding Initiative* 5 (junio 2013). 31 de enero de 2014. <<http://jtei.revues.org/788>>.
- Fradejas Rueda, José Manuel. "La codificación XML/TEI de textos medievales". *Memorabilia* 12 (2009-2010): 219-47. 31 de enero de 2014. <<http://parnaseo.uv.es/Memorabilia/Memorabilia12/PDFs/Codificacion.pdf>>.
- González-Blanco García, Elena. "Actualidad de las Humanidades Digitales y un ejemplo de ensamblaje poético en la red: ReMetCa". *Cuadernos Hispanoamericanos* 761 (2013): 53-67.
- Hipola, Pedro, y Eito Brun. "Edición digital: formatos y alternativas". *El profesional de la información* 9.10 (octubre 2000): 4-15.
- Manovich, Lev. *Software Takes Command*. Norkfolk: Bloomsbury, 2013. 31 de enero de 2014. <www.manovich.net>.
- Martín de Santa Olalla Sánchez, Aurora. "Una propuesta de codificación morfosintáctica para corpus de referencia en lengua española". *Estudios de Lingüística del Español* 3 (1999). 31 de enero de 2014. <<http://elies.rediris.es/elies3/>>.
- Morrás, María. "Las ediciones de textos medievales en la web". Comunicación presentada en el *Congreso Humanidades en la Red, mundo medieval*, celebrado en Barcelona del 23 al 25 de febrero de 2011.

- Negroponte, Nicholas. *El mundo digital*. Madrid: Grupo Zeta, 1995.
- Spence, Paul. “Teatro clásico y humanidades digitales: el cruce entre método, proceso y nuevas tecnologías”. *Teatro de Palabras* 7 (2013): 9-38. 17 de abril de 2014. <https://kclpure.kcl.ac.uk/portal/files/10298890/Paul_Spence_Teatro_Clasico_y_Humanidades_Digitales.pdf>.
- Vandenbussche, Pierre-Yves, y Bernard Vatant. *Metadata Recommendations for Linked Open Data Vocabularies*. Version 1.1. 19 agosto 2012. 24 de abril de 2014. <http://lov.okfn.org/Recommendations_Vocabulary_Design.pdf>.
- Zanoni, Leandro. *El imperio digital*. Buenos Aires: Ediciones B Argentina, 2008.